```
   1 TCCTCCGGTC GCCCGCCCTC GGGGCAGCTA GTGGCGCAGC CCCCCGCCCG
  51 CGGCCCTGGC CTCCCGGGCG GCGCGGCAGG GGAGGGGTTA AGCTGCCGCA
 101 GGGACCGCCG CGTGCGGGGC GAGAGGGAGC CCCCGGTGGG GGTGGCGCAG
 151 CCGGCGGGGT TCGGTCCGAG CCCGGTGGGA GGCTCCCGGA GCGCAGCCTG
 201 GGCCCAGCCC ACCCCGCGCC GGCGGCCATG GCAGGCACCC TGGACCTGGA
 251 CAAGGGCTGC ACGGTGGAGG AGCTGCTCCG CGGGTGCATC GAAGCCTTCG
 301 ATGACTCCGG GAAGGTGCGG GACCCGCAGC TGGTGCGCAT ATTCCTCATG
 351 ATGCACCCCT GGTACATCCC CTCCTCTCAG CTGGCGGCCA AGCTGCTCCA
 401 CATCTACCAA CAATCCCGGA AGGACAACTC CAATTCCCTG CAGGTGAAAA
 451 CGTGCCACCT GGTCAGGTAC TGGATCTCCG CCTTCCCAGC GGAGTTTGAC
 501 TTGAACCCGG AGTTGGCTGA GCAGATCAAG GAGCTGAAGG CTCTGCTAGA
 551 CCAAGAAGGG AACCGACGGC ACAGCAGCCT AATCGACATA GACAGCGTCC
 601 CTACCTACAA GTGGAAGCGG CAGGTGACTC AGCGGAACCC TGTGGGACAG
 651 AAAAAGCGCA AGATGTCCCT GTTGTTTGAC CACCTGGAGC CCATGGAGCT
 701 GGCGGAGCAT CTCACCTACT TGGAGTATCG CTCCTTCTGC AAGATCCTGT
 751 TTCAGGACTA TCACAGTTTC GTGACTCATG GCTGCACTGT GGACAACCCC
 801 GTCCTGGAGC GGTTCATCTC CCTCTTCAAC AGCGTCTCAC AGTGGGTGCA
 851 GCTCATGATC CTCAGCAAAC CCACAGCCCC GCAGCGGGCC CTGGTCATCA
 901 CACACTTTGT CCACGTGGCG GAGAAGCTGC TACAGCTGCA GAACTTCAAC
 951 ACGCTGATGG CAGTGGTCGG GGGCCTGAGC CACAGCTCCA TCTCCCGCCT
1001 CAAGGAGACC CACAGCCACG TTAGCCCTGA GACCATCAAG CTCTGGGAGG
1051 GTCTCACGGA ACTAGTGACG GCGACAGGCA ACTATGGCAA CTACCGGCGT
1101 CGGCTGGCAG CCTGTGTGGG CTTCCGCTTC CCGATCCTGG GTGTGCACCT
1151 CAAGGACCTG GTGGCCCTGC AGCTGGCACT GCCTGACTGG CTGGACCCAG
1201 CCCGGACCCG GCTCAACGGG GCCAAGATGA AGCAGCTCTT TAGCATCCTG
1251 GAGGAGCTGG CCATGGTGAC CAGCCTGCGG CCACCAGTAC AGGCCAACCC
1301 CGACCTGCTG AGCCTGCTCA CGGTGTCTCT GGATCAGTAT CAGACGGAGG
1351 ATGAGCTGTA CCAGCTGTCC CTGCAGCGGG AGCCGCGCTC CAAGTCCTCG
1401 CCAACCAGCC CCACGAGTTG CACCCCACCA CCCCGGCCCC CGGTACTGGA
1451 GGAGTGGACC TCGGCTGCCA AACCCAAGCT GGATCAGGCC CTCGTGGTGG
1501 AGCACATCGA GAAGATGGTG GAGTCTGTGT TCCGGAACTT TGACGTCGAT
1551 GGGGATGGCC ACATCTCACA GGAAGAATTC CAGATCATCC GTGGGAACTT
1601 CCCTTACCTC AGCGCCTTTG GGGACCTCGA CCAGAACCAG GATGGCTGCA
1651 TCAGCAGGGA GGAGATGGTT TCCTATTTCC TGCGCTCCAG CTCTGTGTTG
1701 GGGGGGCGCA TGGGCTTCGT ACACAACTTC CAGGAGAGCA ACTCCTTGCG
1751 CCCCGTCGCC TGCCGCCACT GCAAAGCCCT GATCCTGGGC ATCTACAAGC
1801 AGGGCCTCAA ATGCCGAGCC TGTGGGAGTGA ACTGCCACAA GCAGTGCAAG
1851 GATCGCCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG TGAGCCTGGA
1901 GGGGTCTGCA CCCTCACCCT CACCCATGCA CAGCCACCAT CACCGCGCCT
1951 TCAGCTTCTC TCTGCCCCGC CCTGGCAGGC GAGGCTCCAG GCCTCCAGCA
2001 ATCCCCCTCC CAGCAGAGAT CCGTGAGGAG GAGGTACAGA CGGTGGAGGA
2051 TGGGGTGTTT GACATCCACT TGTAATAGAT GCTGTGGTTG GATCAAGGAC
2101 TCATTCCTGC CTTGGAGAAA ATACTTCAAC CAGAGCAGGG AGCCTGGGGG
2151 TGTCGGGGCA GGAGGCTGGG GATGGGGGTG GGATATGAGG GTGGCATGCA
2201 GCTGAGGGCA GGGCCAGGGC TGGTGTCCCT AAGGTTGTAC AGACTCTTGT
2251 GAATATTTGT ATTTTCCAGA TGGAATAAAA AGGCCCGTGT AATTAAAAAA
2301 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA      (SEQ ID NO:1)
```

FEATURES:
5'UTR: 1-227
Start Codon: 228
Stop Codon: 2073
3'UTR: 2076

FIGURE 1, page 1 of 2

Homologous proteins:

<u>Top 10 BLAST Hits</u>

	Score	E
CRA|1000682340958 /altid=gi|6358505 /def=gb|AAF07219.1|AF043722...	1293	0.0
CRA|18000005086608 /altid=gi|5031623 /def=ref|NP_005816.1| RAS ...	1241	0.0
CRA|18000005188697 /altid=gi|6755290 /def=ref|NP_035372.1| RAS,...	1202	0.0
CRA|18000005205935 /altid=gi|7662334 /def=ref|NP_056191.1| KIAA...	618	e-175
CRA|18000005188699 /altid=gi|3928857 /def=gb|AAC79700.1| (AF081...	533	e-150
CRA|18000005152782 /altid=gi|9507035 /def=ref|NP_062084.1| RAS ...	531	e-149
CRA|18000005192860 /altid=gi|7242201 /def=ref|NP_035376.1| RAS ...	529	e-149
CRA|18000005192861 /altid=gi|4038292 /def=gb|AAC97349.1| (AF106...	526	e-148
CRA|18000005188698 /altid=gi|5032025 /def=ref|NP_005730.1| RAS ...	525	e-148
CRA|1000733831533 /altid=gi|6650545 /def=gb|AAF21898.1|AF081197...	525	e-148

<u>BLAST dbEST hits:</u>

	Score	E
gi|5432583 /dataset=dbest /taxon=9606 ...	1310	0.0
gi|9876673 /dataset=dbest /taxon=960...	1281	0.0
gi|11286864 /dataset=dbest /taxon=96...	1249	0.0
gi|11285315 /dataset=dbest /taxon=96...	1207	0.0
gi|5432584 /dataset=dbest /taxon=9606 ...	733	0.0
gi|4372300 /dataset=dbest /taxon=9606 ...	720	0.0
gi|12295751 /dataset=dbest /taxon=96...	700	0.0
gi|12288965 /dataset=dbest /taxon=96...	599	e-168
gi|6920402 /dataset=dbest /taxon=960...	573	e-161
gi|2005039 /dataset=dbest /taxon=9606 ...	573	e-161

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

<u>From BLAST dbEST hits:</u>

gi\|5432583   Testis
gi\|9876673   Liver-non-cancerous
gi\|11286864 Brain glioblastoma
gi\|11285315 Brain glioblastoma
gi\|5432584   Testis
gi\|4372300   B Cell Chronic lymphatic leukemia
gi\|12295751 Adult marrow
gi\|12288965 Adult marrow
gi\|6920402   Lymph germinal center B cell
gi\|2005039   Lymph

<u>From tissue screening panels:</u>

Leukocyte

```
  1 MAGTLDLDKG CTVEELLRGC IEAFDDSGKV RDPQLVRIFL MMHPWYIPSS
 51 QLAAKLLHIY QQSRKDNSNS LQVKTCHLVR YWISAFPAEF DLNPELAEQI
101 KELKALLDQE GNRRHSSLID IDSVPTYKWK RQVTQRNPVG QKKRKMSLLF
151 DHLEPMELAE HLTYLEYRSF CKILFQDYHS FVTHGCTVDN PVLERFISLF
201 NSVSQWVQLM ILSKPTAPQR ALVITHFVHV AEKLLQLQNF NTLMAVVGGL
251 SHSSISRLKE THSHVSPETI KLWEGLTELV TATGNYGNYR RRLAACVGFR
301 FPILGVHLKD LVALQLALPD WLDPARTRLN GAKMKQLFSI LEELAMVTSL
351 RPPVQANPDL LSLLTVSLDQ YQTEDELYQL SLQREPRSKS SPTSPTSCTP
401 PPRPPVLEEW TSAAKPKLDQ ALVVEHIEKM VESVFRNFDV DGDGHISQEE
451 FQIIRGNFPY LSAFGDLDQN QDGCISREEM VSYFLRSSSV LGGRMGFVHN
501 FQESNSLRPV ACRHCKALIL GIYKQGLKCR ACGVNCHKQC KDRLSVECRR
551 RAQSVSLEGS APSPSPMHSH HHRAFSFSLP RPGRRGSRPP AIPLPAEIRE
601 EEVQTVEDGV FDIHL  (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3
```
        1      113-116 RRHS
        2      144-147 RKMS
        3      584-587 RRGS
```

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 7
```
        1       27-29  SGK
        2       63-65  SRK
        3      126-128 TYK
        4      134-136 TQR
        5      269-271 TIK
        6      349-351 SLR
        7      506-508 SLR
```

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 9
```
        1       12-15  TVEE
        2       63-66  SRKD
        3      117-120 SLID
        4      163-166 TYLE
        5      339-342 SILE
        6      373-376 TEDE
        7      447-450 SQEE
        8      476-479 SREE
        9      605-608 TVED
```

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 4
```
        1       19-24  GCIEAF
        2      249-254 GLSHSS
        3      284-289 GNYGNY
        4      492-497 GGRMGF
```

```
[5] PDOC00009 PS00009 AMIDATION
Amidation site

        582-585 PGRR

[6] PDOC00018 PS00018 EF_HAND
EF-hand calcium-binding domain

Number of matches: 2
        1       439-451 DVDGDGHISQEEF
        2       468-480 DQNQDGCISREEM

[7] PDOC00379 PS00479 DAG_PE_BIND_DOM_1
Phorbol esters / diacylglycerol binding domain

        499-548 HNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVEC
```

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	34	54	0.713	Putative
2	195	215	0.653	Putative
3	238	258	0.788	Putative

FIGURE 2, page 2 of 7

BLAST Alignment to Top Hit:

```
>CRA|18000005086608 /altid=gi|5031623 /def=ref|NP_005816.1| RAS guanyl
           releasing protein 2 (calcium and DAG-regulated); calcium
           and diacylglycerol-regulated guanine nucleotide exchange
           factor I [Homo sapiens] /org=Homo sapiens /taxon=9606
           /dataset=nraa /length=609
        Length = 609
 Score = 1241 bits (3176), Expect = 0.0
 Identities = 608/615 (98%), Positives = 609/615 (98%)
 Frame = +3


Query: 228   MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407
             MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIY
Sbjct: 1     MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIY 60


Query: 408   QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587
             QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID
Sbjct: 61    QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 120


Query: 588   IDSVPTYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767
             IDSVPTYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121   IDSVPTYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180


Query: 768   FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 947
             FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF
Sbjct: 181   FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 240


Query: 948   NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFR 1127
             NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFR
Sbjct: 241   NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFR 300


Query: 1128  FPILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL 1307
             FPILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL
Sbjct: 301   FPILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL 360


Query: 1308  LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 1487
             LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ
Sbjct: 361   LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 420


Query: 1488  ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 1667
             ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM
Sbjct: 421   ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 480


Query: 1668  VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
             VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481   VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540


Query: 1848  KDRLSVECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPLPAEIRE 2027
             KDRLSVECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP      EIRE
Sbjct: 541   KDRLSVECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP------EIRE 594


Query: 2028  EEVQTVEDGVFDIHL 2072
             EEVQTVEDGVFDIHL
Sbjct: 595   EEVQTVEDGVFDIHL 609   (SEQ ID NO:4)
```

```
>CRA|1000682340958 /altid=gi|6358505 /def=gb|AAF07219.1|AF043722_1
         (AF043722) guanine exchange factor MCG7 isoform 1 [Homo
         sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa
         /length=671
          Length = 671
  Score = 1293 bits (3309), Expect = 0.0
  Identities = 639/670 (95%), Positives = 643/670 (95%), Gaps = 4/670 (0%)
  Frame = +3

Query: 75    GRGGVKLPQGPPRAGREGAPGGGGAAG----GVRSEPGGRLPERSLGPAHPAPAAMAGTL 242
             GRG  P  + +E   G   +G    GVRSEPGGRLPERSLGPAHPAPAAMAGTL
Sbjct: 8     GRGTQGWPGSSEQHVQEATSSAGLHSGVDELGVRSEPGGRLPERSLGPAHPAPAAMAGTL 67

Query: 243   DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRK 422
             DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIYQQSRK
Sbjct: 68    DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRK 127

Query: 423   DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP 602
             DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP
Sbjct: 128   DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP 187

Query: 603   TYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 782
             TYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG
Sbjct: 188   TYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 247

Query: 783   CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA 962
             CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA
Sbjct: 248   CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA 307

Query: 963   VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG 1142
             VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG
Sbjct: 308   VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG 367

Query: 1143  VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 1322
             VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT
Sbjct: 368   VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 427

Query: 1323  VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQALVVE 1502
             VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQALVVE
Sbjct: 428   VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQALVVE 487

Query: 1503  HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMVSYFL 1682
             HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMVSYFL
Sbjct: 488   HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMVSYFL 547

Query: 1683  RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 1862
             RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS
Sbjct: 548   RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 607

Query: 1863  VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPLPAEIREEEVQT 2042
             VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP      EIREEEVQT
Sbjct: 608   VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP------EIREEEVQT 661

Query: 2043  VEDGVFDIHL 2072
             VEDGVFDIHL
Sbjct: 662   VEDGVFDIHL 671   (SEQ ID NO:5)
```

```
>CRA|18000005188697 /altid=gi|6755290 /def=ref|NP_035372.1| RAS,
         guanyl releasing protein 2; RAP 1A protein-specific
         guanine nucleotide exchange factor 1; CalDAG-GEFI [Mus
         musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
         /length=608
            Length = 608
 Score = 1202 bits (3076), Expect = 0.0
 Identities = 589/615 (95%), Positives = 597/615 (96%)
 Frame = +3

Query: 228   MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407
             MA TLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLA+KLLH Y
Sbjct: 1     MASTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLASKLLHFY 60

Query: 408   QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587
             QQSRKDNSNSLQVKTCHLVRYW+SAFPAEFDLNPELAE IKELKALLDQEGNRRHSSLID
Sbjct: 61    QQSRKDNSNSLQVKTCHLVRYWVSAFPAEFDLNPELAEPIKELKALLDQEGNRRHSSLID 120

Query: 588   IDSVPTYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767
             I+SVPTYKWKRQVTQRNPV QKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121   IESVPTYKWKRQVTQRNPVEQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768   FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 947
             FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTA QRALVITHFVHVAEKLLQLQNF
Sbjct: 181   FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTATQRALVITHFVHVAEKLLQLQNF 240

Query: 948   NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFR 1127
             NTLMAVVGGLSHSSISRLKETHSHVSP+TIKLWEGLTELVTATGNY NYRRRLAACVGFR
Sbjct: 241   NTLMAVVGGLSHSSISRLKETHSHVSPDTIKLWEGLTELVTATGNYSNYRRRLAACVGFR 300

Query: 1128  FPILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL 1307
             FPILGVHLKDLVALQLALPDWLDP RTRLNGAKM+QLFSILEELAMVTSLRPPVQANPDL
Sbjct: 301   FPILGVHLKDLVALQLALPDWLDPGRTRLNGAKMRQLFSILEELAMVTSLRPPVQANPDL 360

Query: 1308  LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 1487
             LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTS AKPKLDQ
Sbjct: 361   LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSVAKPKLDQ 420

Query: 1488  ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 1667
             ALV EHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM
Sbjct: 421   ALVAEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 480

Query: 1668  VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
             +SYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481   ISYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848  KDRLSVECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPLPAEIRE 2027
             KDRLSVECRRRAQSVSLEGSAPSPSP H+ HHRAFSFSLPRPGRR SRPP      EIRE
Sbjct: 541   KDRLSVECRRRAQSVSLEGSAPSPSPTHT-HHRAFSFSLPRPGRRSSRPP------EIRE 593

Query: 2028  EEVQTVEDGVFDIHL 2072
             EEVQTVEDGVFDIHL
Sbjct: 594   EEVQTVEDGVFDIHL 608   (SEQ ID NO:6)
```

>CRA|18000005205935 /altid=gi|7662334 /def=ref|NP_056191.1| KIAA0846
          protein [Homo sapiens] /org=Homo sapiens /taxon=9606
          /dataset=nraa /length=689
          Length = 689
 Score = 618 bits (1576), Expect = e-175
 Identities = 314/597 (52%), Positives = 409/597 (67%), Gaps = 5/597 (0%)
 Frame = +3

Query: 234   GTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQ 413
             G+ L K T++ELL CIE FDD+G++ + L RI L+MH WY+ S++LA KLL +Y+
Sbjct: 2   GSSGLGKAATLDELLCTCIEMFDDNGELDNSYLPRIVLLMHRWYLSSTELAEKLLCMYRN 61

Query: 414   SRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDID 593
             + ++ N ++K C+ +RYWI FPAEF+L+ L   +E + + Q G +H SLIDI
Sbjct: 62   ATGESCNEFRLKICYFMRYWILKFPAEFNLDLGLIRMTEEFREVASQLGYEKHVSLIDIS 121

Query: 594   SVPTYKWKRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFV 773
             S+P+Y W R+VTQR V KK K LLFDHLEP+ELAEHLT+LE++SF +I F DY S+V
Sbjct: 122   SIPSYDWMRRVTQRKKVS-KKGKACLLFDHLEPIELAEHLTFLEHKSFRRISFTDYQSYV 180

Query: 774   THGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNT 953
              HGC +NP LER I+LFN +S+WVQLM+LSKPT QRA VIT F++VA+KLLQL+NFNT
Sbjct: 181   IHGCLENNPTLERSIALFNGISKWVQLMVLSKPTPQQRAEVITKFINVAKKLLQLKNFNT 240

Query: 954   LMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFP 1133
             LMAVVGGLSHSSISRLKETHSH+S E K W +TELV++ GNY NYR+ A C GF+ P
Sbjct: 241   LMAVVGGLSHSSISRLKETHSHLSSEVTKNWNEMTELVSSNGNYCNYRKAFADCDGFKIP 300

Query: 1134 ILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLS 1313
             ILGVHLKDL+A+ +   PDW +   ++N KM QL L EL + +   ++ N DL++
Sbjct: 301   ILGVHLKDLIAVHVIFPDWTE--ENKVNIVKMHQLSVTLSELVSLQNASHHLEPNMDLIN 358

Query: 1314 LLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQAL 1493
             LLT+SLD Y TED++Y+LSL EPR+ SPTSPT+   P +P V EW   PK D +
Sbjct: 359   LLTLSLDLYHTEDDIYKLSLVLEPRNSKSPTSPTT---PNKPVVPLEWALGVMPKPDPTV 415

Query: 1494 VVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMVS 1673
             + +HI K+VESVFRN+D D DG+ISQE+F+ I NFP+L +F LD++QDG IS++EM++
Sbjct: 416   INKHIRKLVESVFRNYDHDHDGYISQEDFESIAANFPFLDSFCVLDKDQDGLISKDEMMA 475

Query: 1674 YFLRSSSVLGGRM--GFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
             YFLR+ S L +M GF+HNFQE L+P C HC   + GI KQG KC+ CG NCHKQC
Sbjct: 476   YFLRAKSQLHCKMGPGFIHNFQEMTYLKPTFCEHCAGFLWGIIKQGYKCKDCGANCHKQC 535

Query: 1848 KDRLSVECRRRAQSVSL---EGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPL 2009
             KD L + CRR A++ SL   GS P   +   F F   G R   AI L
Sbjct: 536   KDLLVLACRRFARAPSLSSGHGSLPGSPSLPPAQDEVFEFPGVTAGHRDLDSRAITL 592   (SEQ
ID NO:7)

>CRA|18000005188699 /altid=gi|3928857 /def=gb|AAC79700.1| (AF081196)
          calcium and DAG-regulated guanine nucleotide exchange
          factor II [Rattus norvegicus] /org=Rattus norvegicus
          /taxon=10116 /dataset=nraa /length=795
          Length = 795
 Score = 533 bits (1358), Expect = e-150
 Identities = 267/590 (45%), Positives = 390/590 (65%), Gaps = 12/590 (2%)
 Frame = +3

Query: 156   GVRSEPGGRLPERSLGPAHPAPAAMAGTLD---------LDKGCTVEELLRGCIEAFDDS 308
             G R+ P GRL +S   PA ++A   L KG ++++L+ CI++FD
Sbjct: 17   GSRAGPKGRLEAKSTNSPLPAQPSLAQITQFRMMVSLGHLAKGASLDDLIDSCIQSFDAD 76

FIGURE 2, page 6 of 7

```
Query:  309  GKV-RDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRKDNSNSLQVKTCHLVRYWISAF 485
             G + R  QL+++ L MH   I S++L  KL+++Y+ + + NS  + +K C+ VRYWI+ F
Sbjct:   77  GNLCRSNQLLQVMLTMHRIIISSAELLQKLMNLYKDALEKNSPGICLKICYFVRYWITEF 136

Query:  486  PAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVPTYKWKRQVTQRNPVG-QKKRK 662
             F ++  L  ++E + L+   G  H  LID  + +  W R++TQR      KKRK
Sbjct:  137  WIMFKMDASLTSTMEEFQDLVKANGEESHCHLIDTTQINSRDWSRKLTQRIKSNTSKKRK 196

Query:  663  MSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHGCTVDNPVLERFISLFNSVSQ 842
             +SLLFDHLEP EL+EHLTYLE++SF +I F DY +++ + C  +NP +ER I+L N +SQ
Sbjct:  197  VSLLFDHLEPEELSEHLTYLEFKSFRRISFSDYQNYLVNSCVKENPTMERSIALCNGISQ 256

Query:  843  WVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMAVVGGLSHSSISRLKETHSHV 1022
             WVQLM+LS+PT  RA V  F+HVA+KL QLQNFNTLMAV+GGL HSSISRLKET SHV
Sbjct:  257  WVQLMVLSRPTPQLRAEVFIKFIHVAQKLHQLQNFNTLMAVIGGLCHSSISRLKETSSHV 316

Query: 1023  SPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILGVHLKDLVALQLALPDWLDPA 1202
             E K+  +TEL+++  NY NYRR   C F+ PILGVHLKDL++L  A+PD+L+
Sbjct:  317  PHEINKVLGEMTELLSSCRNYDNYRRAYGECTHFKIPILGVHLKDLISLYEAMPDYLEDG 376

Query: 1203  RTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLTVSLDQYQTEDELYQLSLQRE 1382
             ++N K+  L++ + EL  + + PP+ AN DL+ LLT+SLD Y TEDE+Y+LS  RE
Sbjct:  377  --KVNVQKLLALYNHINELVQLQDVAPPLDANKDLVHLLTLSLDLYYTEDEIYELSYARE 434

Query: 1383  PRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQALVVEHIEKMVESVFRNFDVGDGH 1562
             PR+ +P        P +PPV+ +W S  PK D   + +H+++MV SVF+N+D+D DG+
Sbjct:  435  PRNHRAPP-----LTPSKPPVVVDWASGVSPKPDPKTISKHVQRMVDSVFKNYDLDQDGY 489

Query: 1563  ISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMVSYFLRSSSVLGG-RMGFVHNFQES 1739
             ISQEEF+ I  +FP+  +F +D+++ +G ISR+E+ +YF+R+SS+     +GF HNFQE+
Sbjct:  490  ISQEEFEKIAASFPF--SFCVMDKDREGLISRDEITAYFMRASSIYSKLGLGFPHNFQET 547

Query: 1740  NSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVECRRRAQS 1889
             L+P  C +C  + G+ KQG +C+ CG+NCHKQCKD +  EC++R++S
Sbjct:  548  TYLKPTFCDNCAGFLWGVIKQGYRCKDCGMNCHKQCKDLVVFECKKRSKS 597  (SEQ ID
NO:8)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00617	RasGEF domain	123.5	4e-33	1
PF00130	Phorbol esters/diacylglycerol binding domain	59.5	3.6e-14	1
PF00036	EF hand	21.8	0.00027	2
PF01237	Oxysterol-binding protein	3.5	4.2	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF01237	1/1	249	272	..	1	24	[.	3.5	4.2
PF00617	1/1	148	336	..	1	227	[]	123.5	4e-33
PF00036	1/2	430	458	..	1	29	[]	17.4	0.0047
PF00036	2/2	463	482	..	5	24	..	6.7	4.9
PF00130	1/1	499	548	..	1	51	[]	59.5	3.6e-14

```
   1 ACAGAAAGGT CCTGTTTCTA AGTCTTACAT TACCAAGACT GAGGTGCGGG
  51 GGCGGTCCTG GATCCCCCGC CCCAAGGCTG GGAGGGGCAC GCCTCGGAAG
 101 GGAGGTTTGG GGTCGGTGGT TTCACAGTGA GTGTGTCTGA AGCCAAATGG
 151 TCGGAAACCG TTACCCGCTC TCCTAGGCCC GGCTAGTGGG GACCCCAACC
 201 GCCTGCGGCT GCCCCTCCCA AGTTCCTCCC TGTTGGCCAG GCATCCAGGT
 251 CTCCAGTCTC CGAGCTGCGG AGAACCCACC GCCACATGCG GCTGCCCCTT
 301 TCCATTCGAC CCTGTGGGGA GCCAGGCTTC CGGGGCCCCG TTCCTCCTGT
 351 GTGAACTGGG CCCCCCGCCC CCATTCCCAG ACATCAAGGC CGCGTCTCCA
 401 GATAGCCACG ATTTCATTCC TCGCTCCCCA CAGGTCCCTC TCCCCAAAAT
 451 ATTCCCATCT TGTCCTAGCC CATCCCCCAG ACTATCTCAA GGACCAGCTG
 501 TCCCCACGCC CCCGACCTCC ACTAGGCCTG TGCCACCCGC TGCCTGCAGG
 551 AAGACGCCCG GTCCCGGGCC GGGTTAGCCC CATGGGAACG GTTTGTCTCG
 601 AAAACAGGAA CCCGAGCTGG GGGCTGGGCG GGGCGCCCCT TCCCCACCGC
 651 AGTCCGCTTC CTGCCCCTCC CGGCTTCCTC CGCCCGACAC CCAGGCAGGG
 701 CGGGGGGCAC TGGGGCGTCC GCGGTTGGGG GAGGGGCTCT TCGTTTCGGT
 751 CCCCCCTCCC GCGTCCCGGG CGGCGGGGCC TCCGGTCGCC CGCCTCGGGG
 801 CAGCTAGTGG CGCAGCCCCC CGCCCGCGGC CCTGGCCTCC CGGGCGGCGC
 851 GGCAGGGGAG GGGTTAAGCT GCCGCAGGGA CCGCCGCGTG CGGGGCGAGA
 901 GGGAGCCCCC GGTGGGGGTG GCGCAGCCGG CGGTGCGGAG CTCCGCGCAG
 951 GGGCGGAGGG GGGAGGGGGC AGCCTGGCGC GGGGGCGGGG GCGGGGCGGC
1001 GGGGAGCGGG GCCGCGGCGT GGAGAGCGGG CGGGAGCCGC AGCCGCAGCG
1051 AGGCCGGCGG GCGGGAGCGC ACGGAGGTGG GGTCGGCCAG GCCGGTGCGG
1101 GCTCCTTGCG GCAGGTCCCA AGAGTGAGTG GGCGAGCGCG GGCGGGGCGC
1151 CAGGCGAAGG AGGGCGCGGC CCCCAGCGAC TCCCCCCCCG CCCAGGGCGG
1201 CGCGGGCGGG CTGGGGGCGG CGAGCGGGTG GGGAGTCTGC GGCCCGGGTC
1251 TGGGAGAGGG GGCAGCGGCC ACGAGAGCTA AGGCGCGCTG GATCCCCGGA
1301 GGGCGGAGGA CCTCCACGGT GCACCCAGCT TTTCCCAGCC ACCTTCCAGC
1351 GGGGCCCTCC CCCGCGTACC CCCATTTGGC AGATGAGAAA ATTGAGGCTC
1401 CCAGAGGCCA AGTGATTCTC AAGGTCACAC GAGGAAGCGG TAGAGCCAGG
1451 CGGGGACGGC TCTGGGTGGC TCTTAGGAAA AGTCCGCCTG AGAACTCCGT
1501 ACAGGAGCTC CCCTGTCCTC CAGCCTGGGG GAGTGAGTAT GTGTAGGGCC
1551 GGGGTACCTT TCCGTGGGGC AAGGCTCTGC CAAAATCTGG GAGTGAGGGG
1601 AGTCAGGGAG CTGGGGCCGC AGGGCGGGCC CTGCACCGCA AATGGGAGGG
1651 GGGCGACGGA ATGGGCGTGC GCACCCATGG GGGTGTGTGC ATGTGTGTGG
1701 GAGTGTACAT GCGTGGAGAG GCACTGCCTT GCGTGTGTGC ACACGTGTGA
1751 GGATGTCAGC GCCTGTGTGG CCGCGGGACT CAAGGCTGGC CTGGCTCAAG
1801 TGAACAGCAC GTCCAGGAGG CGACCTCGTC CGCGGGTTTG CATTCTGGGG
1851 TGGACGAGCT GGGTATGTGT GCCTGAGGGT TTCTTCGTGC AGGTGTGCAC
1901 AGGGTGTGGG TGCCATTGTG TGTGAGAGAC GGAGGATGGG GAGGCCGGTG
1951 CCTGTGGCCC GGTGCGTGTA AGTGCGGACG CCTGCACCTC CACTTAGGTC
2001 CCCGGCCTCC GACGACTAAC TTGGGTGTGG AGTGTTTGCC CCTGCCAGGG
2051 TGCGTATGAC CCCGCCAGTG ACCGGAGTTG CTAATGGTGT CATGCACCCA
2101 CCGGCCACCC TTGGCGCGAG CGCCCCCCTC TGGACACCCT GCTCCGTGCG
2151 CGCTCACAGT TCGCCTGTGC GGGGCCGGGG CCAGGGTCAG GAGCCGGGGA
2201 TAGGGAGGAA GAGGGCCTGT GGACAAGCTG AGCCGGGACC CCTGGGACCT
2251 TTGCGGAGGT GGCCTGGGAG CGCTCAGTTC CCAGGCTGAG GCTTCCCGCT
2301 GACGCCTCCT GGCCGCAGCG GGCTCCCCCC GCCCCAGGAA TGTTCCTCTC
2351 CCATCCAGTC CGCCTCCCCT AGGGCAGGCC CCCTGGGGGC TGCCGCAGCC
2401 CCGCCTCGCC TTCCTGGGCT CCCGGGAGGG GGCGAGGCGA GCAGGACGCC
2451 TGGGTTCTCT CCCCCCACCT CCCATACCAG GGAGAAATTC CTCCGAGGTC
2501 CCCTCAGGCT CTGGGTTCCC AAAATAACCC TGCGGGGGAA GGGAGGCTGT
2551 GGAGGGAGGG AAGCGGGAGG GGCGCAGAGC CGAGCTGCGG GGTGCTGCAG
2601 GTGCCTCTGG GGAGAGGGCG CGAGGAGAAG GCGCCCTGCG GGGGGCTGGG
2651 CGCCAGCCAG TCCTGGGATC TTGGTTCGTC CCCATCCTCG TGAAGCCCCT
2701 CGGCCTTCCC GCGACTCCGA GGGTGGGCCG GAAGCCTCTC TGCGGGTCCG
2751 TTTCCCAACT GGCGGGTTGC ACCATCCCGG GCCAGACCGT TTAACCCCGG
2801 GAGTGGCCGC GGGGGACAAC TCCGCCCCTG TCCAGCAGGG GGCGTGCCCG
2851 CCCCGCCCCG TTTCTGCCCG CGGGGCCGCT CCCCCGCCCG CGACTCCGCA
2901 GACTCCCGCT CTGCCTCTCC CGGGACAGGG GTTCGGTCCG AGCCCGGTGG
2951 GAGGCTCCCG GAGCGCAGCC TGGGCCCAGC CCACCCCGCG CCGGCGGCCA
3001 TGGCAGGCAC CCTGGACCTG GACAAGGGCT GCACGGTGGA GGAGCTGCTC
3051 CGCGGGTGCA TCGAAGCCTT CGGTGAGTGG CTCGGGAGGG CACACGGAGC
3101 CTGAGCCTAG CCCCGAGTCT GAGCCCGGGT CCCTGCCTCC CAGGCACAGT
```

FIGURE 3, page 1 of 12

```
3151 CCAGGGCACA GCCCTGACCC GGACCCACCC TGCTCCGCAG CGTGCAGTCT
3201 CTTTAACGAA AGCCTCCTCC GCAACGCAGG GCAGAGAGAT GCACGCCCTT
3251 CAGACAGATG AGGTTTCCCT TCTCTAGCCT TCCCCAGCGG CGGCGAAGGG
3301 AGGGCCGGGT CCCGGACTCT GACACTTGAG GGGCATTATC TGTCTCCCGG
3351 GGAATCCGGA GGAACTCGCT ATCTCCGGCC TGGGAGCTGT TTCCGGCTAA
3401 TGGGGGGCGG CTTATCTGGT GAAGGGGTGC CCCTTCCCCC CAAGCGCTCA
3451 GGAAATGACC TCTGGATTCT TGACCCCGGG GAACCCAGGC TCCTTCCGCC
3501 CCAGCTGGTT CCCCTCCGGA CGATGGGCGG CTCGGGCGCT CCCCTCCTCC
3551 AGTCCTCAGG GCGTGCCTAT CTCTCGCCCA CCACACCTTT CCTCTCTAAT
3601 TTGCCTCCTG CTCTCGGAGT CCTGGGCAAG CAGGAGGTGG GCGGGGTCGA
3651 GCGTGCACCC GAAGGACCGA TACCTGGCGG GTTGCGGGGT GAGGATGAGG
3701 CATGGTAGCT GCGGACCCAG CTCAGCCACC TGTCTTTGAC CCTTCGGAGT
3751 CAGATGACTC CGGGAAGGTG CGGGACCCGC AGCTGGTGCG CATGTTCCTC
3801 ATGATGCACC CCTGGTACAT CCCCTCCTCT CAGCTGGCGG CCAAGCTGCT
3851 CCACATATAT CCTTCGCCGG CCTTGCCAAG GCCCCGCCG TCGGAGCCCA
3901 TGCGCAGCCC CTCTGCCCAG CCCAGGTGCA GAATGAGCCT CGCTCCTAAG
3951 TATAGGCCAC TCCTTATCCC AGAGCTCAGG CGTCGTCCCA GCCTCCAACT
4001 AGGGCCTAGG CTCTGCCCCC TCCTTGCTCC TAGCGACTCG GTCCTGTCCC
4051 CAGGCTCTGT CCCCAGCCGA GGCCTTGCCC TCCTTCTCCC TAGAGTCTAG
4101 GGCCTGCCCC TGCTTCAGGC TTGGGTGCGC CCCGTGCATC TCTCTCTCCC
4151 AGAGCCCAGG CTTTGCTTTC AGCCTCCCTC AGCACCTAGT CCTCCACCCC
4201 CACCTCCAAC CCCTCCCAGA GCTCAAGCCT CACCCCCAGC ATCTCCGCAG
4251 AGCGCAAGCC CCATCCCTAG AACGTGTCTC CTAGAACCAG GCCCCGCCCC
4301 CAGCCTCCCT CCACGCAGGC CTCCCTTTCT AGAGTTAAGC GGCCTCCTTA
4351 ACCCTCTCCT TCACCTACCA ACAATCCCGG AAGGACAACT CCAATTCCCT
4401 GCAGGTGAAA ACGTGCCACC TGGTCAGGTG AGTCTTTCCC CTGGGGCTCT
4451 AGCCCCTCCC CTTTCTCCCT TCTCTCTGGC TTCAGGCTGG CCTGGAGGAG
4501 GGGGCAGGGC GCTGTTTCTG GGAGTGGGTT TGAACCCTGG CTTGTCCGGG
4551 TGGGCAGTGC TGCCACAGGC TCACCCCTTC CTGGGTCTGG GCCTTAATTT
4601 TCTTTTCTGC GCAGTGCGGG TGGTTGTCTC AAGGGTCTAA TGTACACTTG
4651 GAGTGGCGAA GGAAAGAGCT GGAACCATAG TTTGAGGGTC TTTTTGCTTA
4701 GGTGACTATA ATCTCAAATA GCTCCTTGCA GCCTGCTGGG TGATGGTGGG
4751 GGAAGGGCTA TCTTGGGTGA CTCCCCGCTC CTCCAGGTAC TGGATCTCCG
4801 CCTTCCCAGC GGAGTTTGAC TTGAACCCGG AGTTGGCTGA GCAGATCAAG
4851 GAGCTGAAGG CTCTGCTAGA CCAAGAAGGG AACCGACGGC ACAGCAGCCT
4901 AATCGACATA GACAGCGTGT GCGTGGGGGG AGCACAGAGG GCTGGGGGGG
4951 CACTCAGTAT CCTATACCAT CTGTGCTTAA TAAATGTCTG TTGAACTGAA
5001 TGAGTGAGGG TCATGTTGCT CTCTCGCTTA AAAACCTTCC ATGGCTCCCT
5051 ATTGCCTTCA ACATGCCTCC TCTGGGCAGC TTGGCGTTCC TGCCTCATCT
5101 TCCACTGCCA CCACCCATCC CACACACCTC CTCCTGTAGC TGCGCTGGGT
5151 CGGCTCCCCG TCGGCTGAGC TCTCGAGTCC TTTCTCATCA TGGTGCTCTG
5201 CTCATATCAT CCCCCTTGCT GCCTCCTCCG TGTTACCAAG ACTCAGTTCA
5251 GGCATGAAGT CTCCGTGGGC TCTGAGGGTT CGGGGCTCTT CCGGGGTAGA
5301 ATTTGTCGTT CCCACCTCTG TTTTCCATGG CACTTTGTAC AGACTCCTGT
5351 ACAAAGACCT CTGTACATGT GTCACGCTGT TTTGTGATCA TGTGTTTCTG
5401 TGTCTGTCTC CCTCAGTAGA CTGTGAGCTC CTCGAGGGCA GGAACCGTGT
5451 CTTACTCATC TCTGTATTCC CAGCGCCTAG CACAGTGCCT GGCACAGAGT
5501 ACGTTGTTCA TAAATGTGTG TTGAGTGCAT GACGGGGTGG GGGGAGATGA
5551 GGAGGAGTTG CTGGGACTGG GAACATTCGT GCCTAGGACA GTGCCTCGCA
5601 TTATGTAGGT TCTCAGTAAG CGTGAATGGT GTGTCTGTGT GAGTGGGGGG
5651 CCACGAGGCA TGCGCATGTC CAGCAAAGGG CTCACTACCC CTGCCCCCCC
5701 AGCCCTACCT ACAAGTGGAA GCGGCAGGTG ACTCAGCGGA ACCCTGTGGG
5751 ACAGAAAAAG CGCAAGATGT CCCTGTTGTT TGACCACCTG GAGCCCATGG
5801 AGCTGGCGGA GCATCTCACC TACTTGGAGT ATCGCTCCTT CTGCAAGATC
5851 CTGGTGCGGC CCGAGGGCTG GGGGGTCAGG GGTCCAATGT GGGCTGGAAG
5901 AGAGTTCTAG GAGGGGCAGG GTCCCTGGCG TAGGCTGGGT CACAGGGTGC
5951 ATCAGGGGTT TCAGTGTAAC CACTGAAGGT CAGCTGGAGG GTGAGGAGTG
6001 GCTATCAGTG AGGGGAGAGG CCGGCAAGGT GCTGAGGCCA CTCCTCATGC
6051 CCCCAGTTTC AGGACTATCA CAGTTTCGTG ACTCATGGCT GCACTGTGGA
6101 CAACCCCGTC CTGGAGCGGT TCATCTCCCT CTTCAACAGC GTCTCACAGT
6151 GGGTGCAGCT CATGATCCTC AGCAAACCCA CAGCCCCGCA GCGGGCCCTG
6201 GTCATCACAC ACTTTGTCCA CGTGGCGGAG GTGCCTGCCC CTCCCTCCCG
6251 GTGTCTCCCA ACCACCCCAC ATGCCAGTCA GGCCAACCCT TCCCTTCCCC
```

```
6301 TAACCCACTG CCTTCTCTCT AGATAAGCTG GGCCAAATTC TGGGCCCACT
6351 CAGTGACTCC CTGCCTCTCC GTCCCCATTT GCCTTCCAGA AGCTGCTACA
6401 GCTGCAGAAC TTCAACACGC TGATGGCAGT GGTCGGGGGC CTGAGCCACA
6451 GCTCCATCTC CCGCCTCAAG GAGACCCACA GCCACGTTAG CCCTGAGACC
6501 ATCAAGGTGC CTGGGACTGG GGAGGGGCCG GTGCTTCCCA GGTCTGTCTT
6551 CACTGGGTCC TCCCAGCAGC ACTGGGGGCT GGGCACAGCT GTCCTCATTT
6601 GATAGATATG GAAATGGAGG CTCAGAGGGG TTAAGTGCTT TTCTCAGTTT
6651 GCACAATGGC AACAGCAGAG TGGGGGCTCA CAGGTCGTCA GGGACCCCAA
6701 AGCTAGTACT TTTTTTTTTT TTTTAAGAC AGGGTCTCTC TCTCTGTTGT
6751 CCAGACTGGA GTTCAGTGGT GCAGTCACAA GCTCACTGCA GCCTTGAATT
6801 CCTGAGCTCA ATCGATCCTC CCACCTCAGC CTCCTGAGTA GCTGGGACTA
6851 CAGGTGTACG CCACCATGCC TAATTTTTGT ATTGTTATTA ATTTTTTTTT
6901 TTTTTTTTTA GAGATGGGGT TTTGCCATGT TGCCCAGACT GGTCTTGAAC
6951 TCCTGGGCTC AAGTGATCCG CCTGCCTTGG CCTCCCAAAG TGCTGAGATT
7001 ATGGCTTGAG CCATTGTGCC TTGCCACTTG TAGTTTCTTC TTTTCTTTCT
7051 CCTTCATTTT TTATTATTTT TGAAGTATTT TGAAGTATTG AGTAACATAC
7101 ATATAGAAAA GTATATAAAA ACATATGAGA CTGGGCGTAG TAGCTCACAC
7151 CTGTAATCCC AGCACTTTGG GAGGCTGAGG TGGGCAGATC ACGTGACATC
7201 AGGAGTTTGA GACCAGCCTG GCCAACAAGG TGGAAACCCA TCTCTACTAA
7251 AATACAAAAA TTAGCCAGGC ATGGTGGCAC GCACCTGGAA TCCAAGCTAC
7301 TTGGGAGGCT GAGGCAGGAG GAGAATTACT TGAACTCAGG AGGCGGAGGT
7351 TGCAGTGAGC CAAGATTGTG CCACTTCACT CCAGCCTGGG CGACAGAGTG
7401 AGACTCCATC TAAAAAAAAA GAAAGTATA TAAAAACATA TGAATAGTTT
7451 AAAGAAAAAT TGTAAAGAAA ACACTGTGTA ACTACTGCCC GGGTTGGGAA
7501 ATAGAACCTT GCCAGGCCCC CAAGCGCCCA GCACTTTAGA GCATAACTCC
7551 CTCCCCACGA CTTTTGCAAT GATGATCTTG CTTTTCTTTA TAGCTTCACC
7601 ATGTAGGTAT GCGGTCCAAA ACAATGTGGG GCTTTTTGTT GTCTGTTTTG
7651 AACTTTCTAT GAATGGAATG TTGTTTGTGT TATTTTATGT CTTGCTTTTT
7701 TCATTCCACA TGGTTCTGAG AGTCTTTTCA TTCTGTCATG TGGAGCAATT
7751 GTTTTTTCAT TTTCATTGCC ATATAATATT TTATTGTACG TCTACCCCAA
7801 TTCATTTATT TATTTATTTT TTTGAGATGG AGTCTGTCTC TGTCATCCAG
7851 GCTGGAGTGC GGTGGCGAGA TCTCATCACT GCAACTTCCG TCTCCTGGGT
7901 TTACGTGATT CTCGTGCCTC AGCCTCCTGA GTAGCTGGGA TTATGGGCTC
7951 GTACCACCAC GTCTGGCTAA TTTTTTGTAG AGACAGGCTT TCACCATGTT
8001 GCCGAGGCTG GTCTTGAACT CCTGAGCTCA GGCAATCCAC CCGCTTTAGC
8051 CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACTGCCCC CAGCCTACCC
8101 CAATTTATGT ATTGATTCTA TTGTTGAATG TTGGGGTTTT TCCTTTTCTT
8151 TTCTTTCTTT CTTTTTCTTT CTTTTTTTCT TTTTTTTGGA GAGGGAGTCT
8201 TGCTCTGTCG CCAGGCTGGA GTGCAGTGAC GCTAATTTGG CTCACTGCAT
8251 CACTGCACCC TCTGCCTCCC GGGTTCAAGC GATTCTCCTG CCTCAGCCTC
8301 CTGAGTAGCT GGGACTACAG GCATGCACCA CCACACCCGG CTAATTTTTG
8351 TATTTTTTTA GTAGAGATGA GGTTTCCACC ATGTTGGCCA AGATGGTCTC
8401 CATCTCTTGA CCTCATGATC CATCTGCCAT GGCCTCCCAA AGTGCTGAGA
8451 TTACAAGTGT GAGCCACCAC GCCCAGCTGG TTTTTCCAGT TTTTGCTGTT
8501 TGGACGGGGT GGCTGAGTAT GTTCTTCCAG GTCATTGTCC TGTGCTGCCT
8551 TGCCTCCCTG AGCCTCTGTT TCTCCTGTTA AATGTTGATG ATTCCCTGCA
8601 TCCAGGCCTG GTTTAGAGGT GTGGTGCTTT TGGCAGTGAG TATTGCCTTG
8651 AATTCATGGC AATGAATTCA ATCCCCAGGG GCTGAGAGAG CCAGTCGTGG
8701 GGGACAGTAA GGGAGGTTTT TACTCTTTCA CCTGTCCCTG ACCCTGACTC
8751 CTCCTCACCC CCTCCTACAT TTCCAGGGCT GAGGTAGGGA GGATAGTTGT
8801 GGGGGTATGA CTCCTCTGTC CTTTGTCCCC AGCTCTGGGA GGGTCTCACG
8851 GAACTAGTGA CGGCGACAGG CAACTATGGC AACTACCGGC GTCGGCTGGC
8901 AGCCTGTGTG GGCTTCCGCT TCCCGATCCT GGGTGTGCAC CTCAAGGACC
8951 TGGTGGCCCT GCAGCTGGCA CTGCCTGACT GGCTGGACCC AGCCCGGACC
9001 CGGCTCAACG GGGCCAAGAT GAAGCAGCTC TTTAGCATCC TGGAGGAGCT
9051 GGCCATGGTG ACCAGCCTGC GGCCACCAGT ACAGGCCAAC CCCGACCTGC
9101 TGAGCCTGCT CACGGTGAGG AGCAGGGGGC AGGGAGGTGG GGAGCTGGGC
9151 ACCAGGGGTT GACAGTTTCC CCAGGTCCTG GCTGTGGGCG TGGCCTGGGG
9201 CTCTGGGTTC TGGCCAAGAA ACTGAGATCT AGCGTGGGCT CTGGGGTTTG
9251 GAGTGGATGC TGAGAAGGGG TCCAGGCTCT GGTTGGGGCT GTGGACTGAG
9301 GTCTGATCTC CAGGCTGGTA TGTGGACTGT GGGCAGTTTG AACTGGGCCT
9351 GGGTCCCGGG TTGAGTTCTG GCAATGGGCT GTGTTCTAGG GCTGGGCCAA
9401 GCTCTGCATT CTGTGGGCAG GGGTGGTTTC TAAGCATGGC CCTGGGCTCG
```

```
 9451 GAGTGAAGTT CTGGGCTTGG CTTTACACTT GGTCTTGGGG TCTAGGGTGG
 9501 GAGTTGGGTT CTGGTTTAGA TCCAGACAAG GTTCTAGACA TTGGGCTGGG
 9551 GCTTAAGTGT TAAGGTTTGG AGTGGATTCT TAGCTGCTTC TGGGCTCTGG
 9601 AGGGGATCAG GGTTGAAATC AGAGCTTCTG GCTGGGTTCC GACCTGGCTT
 9651 CTTCCCTGAC ATCTTGGCAA TATGTTGTGT TCAAGGTTTG GGGCCATGCT
 9701 GTGGTTTGAT CTGTGCGCTG GGATGACATG GGGGTTGCTG TGCTGTGTTC
 9751 TAAGCCAGGC TTTGTCCTGA GTCTAGCTTC TGACCCGAGC TCTGGCTGAG
 9801 CTGTGGCCTC TAGGTCGACC TTTGGCCCTG GGCTCTGTGG CCGTGGGCAG
 9851 GGGCCAGTGG GGGTGATCAG ATCTGTGTGT CCCAGGTGTC TCTGGATCAG
 9901 TATCAGACGG AGGATGAGCT GTACCAGCTG TCCCTGCAGC GGGAGCCGCG
 9951 CTCCAAGTCC TCGGTGAGGG GGTACTCCCT CCTCTCCACT CTGCCCTTCC
10001 CTCCTGAGAA TCCCAGGATG TGAGGATGGG AAGAGCTCTT AGCAGCCACC
10051 TCACCCATCC ATCTTGTAGG ACAGAGGCAT CCTGGGGGTA GGGCAGTAGT
10101 GTTGGGCAGA CTTCCCTCTC CCAGGGATTC CCCTCTCTGT TCCCCGGGGC
10151 TCTGGGCTCC CCCTGCCTCT GGCCCTAGCT CAGGCCCGAC CATTTCCATA
10201 GCCAACCAGC CCCACGAGTT GCACCCCACC ACCCCGGCCC CCGGTACTGG
10251 AGGAGTGGAC CTCGGCTGCC AAACCCAAGC TGGATCAGGC CCTCGTGGTG
10301 GAGCACATCG AGAAGATGGT GGAGGTGAGC TCCTGCGGAG CCTGAGCAGT
10351 GTGTGGGGAG AGGCCAGTTT GCCGGAGCAC TGCCCTGGAA GCCAGCACGA
10401 GTGTCCTGTT CAAGACCCAG CACTCAGCCC CTAGGAGTCA CAGGGCCTGG
10451 CAGGCCAGCT GCACGGGGCT GAAGTGCCCC TGGGTAGGGT GGGGGTGGAG
10501 GTATGGAACG GGGGTGGTGT CAGAGACCTC TCTGAGACAC ACCTCATCAA
10551 ATGGACTGGG AACGTGGGAA GGGACAGGAC CTGATGTCCC CTTTACTCTC
10601 CCCTCTTCTG GCTCTGCGTG TCCCTCTGCG TGCCCCAGTC TGTGTTCCGG
10651 AACTTTGACG TCGATGGGGA TGGCCACATC TCACAGGAAG AATTCCAGAT
10701 CATCCGTGGG AACTTCCCTT ACCTCAGCGC CTTTGGGGAC CTCGACCAGA
10751 ACCAGTGAGG AGGGCTGGGG ACCTGGGGGA GAGGGAAGGC AACTCAGCCC
10801 ACTTCTGCCT GGGCTTCAGT TTCTTGTGTG CAAGATGAGG TCACTGAGCC
10851 AGATGATCTT GGCCTGGGAA GCTGCCAGTG TGGGAAAGGG CACTTGCTTT
10901 TGTGGGGAGG AGAGGCTGCC AGCTGTGGAG GCGCAGTGGT ATCTCACAAA
10951 TTCAGACAGA TGGGGGGCTC CACCTGAGTC TTGCAAAGAC TGTGACCTGG
11001 GGACTGTGGC TACAAAAGTG CTGTTTTATT TGTGGAGCTC ACAGCTGTCA
11051 AGAAGTGTGG GCAACTTGAG CTCCTGGATA GTCTGTTCTA ATGAATAGAT
11101 AAGAAAGGTT TGTAATTAGC AGTACCCAGT TGTTTATCAA CAGTTCATAT
11151 GCTGACAATT TGGAAAAACA GCTGGTTCTC TGAAGTAGGT TAAACATGCC
11201 CCCTGAAGCC AGATTCATGC CCTATTTTTG CTGAGCAGAA AAAACTCCAT
11251 TCAAAATTTA AAGTCCATCT CAGGTCGATT TATTTTTTAA TGTTACCTGT
11301 ATTTCAAAAA TCTGTTGTTT TTTATTTCCA CATTACAAAA ATCCACGGTA
11351 AAATAAAATC TAGGTGGTAA AATAAATTTA TAGTGAACAA AATGTTTAAA
11401 GTAAGAAGTG AGAGGCCAGG TGCGGTGCCT CACGCCTGTA ATCCTAGCAC
11451 TTTGGGAGAC TGAGTTGGCA GGATCAATTC AGGCCAGGAG TTTGAGCCCA
11501 GCCTGGGCAA CAGAGTAAGA CCCTGTCTCT ACAAAAATTA TTATTATTAT
11551 TTTTGAGACA GAGTCTCACT CTGTTGCCCA GGCTGGAGTG CAGTGGTACA
11601 ATCTCGGCTC GCTGCAACCT CCACTTCCTG GGTTCAAGTG ATTCTCCTGC
11651 TTCAGCTTCC TGAGTAGCTG GGATTACAGG CATGCATCAC CGTGCCTGGC
11701 TAATTTTTGT ATTTTTAGCA GAGATGGGGT TTTACCATGT TGGCCAGGCT
11751 GGTCTCAAAC TCTTGACCTC AAGTGATCTA CCTGCCTTGG CCCCCCAAAG
11801 TGCTAGGATT ACAGGCATGA GCTACTGCTC CTAGCCTAAA AAAATTTTTT
11851 TTGGGCATGG GTGGCACGTG CCTGTAGTCC CAGCTACTCA GGAGGCTGAG
11901 GCAGGAGGAA CCCTTGAGCC CAGGAGGTTG AGACTGCAGT GAGCTGTCAT
11951 CACACCACTG CACTTCAGCC TGGGTGACTG CGCGAGATCA CCCCCATCAA
12001 AAAAAAAAAA AAAGAAAAA AAAAGGAAGA AATGAAAGTC CCCTCTTTCC
12051 TTTTCCACTG GTAGAAGTTG CCATGATTAA GCACTGTTAA CAATATTAAG
12101 CTTGGCAGTA TGTGGATTCT TCCAGTCTTC TTTTCCCAGG CAGGTGCACA
12151 TTGATAGAGA TTTTGTTTGT TTGGTGTCTG TTTCATGGAC AAACAGGATT
12201 AGAGCATAAA TCTAGTTCTG CTTGTGGCTT TTATCATAGC TGCTTTATTT
12251 CTTCTCCCAG ATTTTAGGCA GAGGTAGTTG AGTTCCATGT TTTCTCCCTG
12301 GGTTGGTGGG TGGATTTTTA TCTAGACCAC CTTTTCAGTG AGAATGACCC
12351 TTTGAGACGA TGGAGGCCTC AGCTTCATGC AGCGGGCTCA GCCTTAACCC
12401 TCCACCTCCT GCAGGCCCCA AGCTGTGTGT GTGTGTGTGT GTGTGTGTGT
12451 GTGTGTGTGT GTGTGTGTGT GTTGGTAAGG GGAAAGCCCC TGGTTGGGTA
12501 TCAAAAACCT AGCACCTGGT TCGGCAGGAG GGAGACCAGC ACCGGCTCCC
12551 CAGGACCAGG CCCAGCTCAC CACTTCATTG TAAAGCTCCC TCTTTGTTTC
```

FIGURE 3, page 4 of 12

```
12601 TGGAACTTGG GTGTTTCCAT TTCTTTCTTA CAAAATTATC TATGCATTTA
12651 CAGCAATTGT TGATATATCT TTAGGCAGCA TCTAGGTACT TGTAGTGGGT
12701 TCTCTTTTTT CTTTTTTCTT TTTTTAATC ACCCTCTCTT TTTTTTGAGA
12751 CAGAGTCTCA CTCTGTCGCT CAGGCTGGAG TGCAATAGCG CGATCTTGGC
12801 TCACTGCAAC CTCTGCCTCC CAGGTTCAAG TAATTCTCAT GCCTCAGCCT
12851 CCCAAGTAGC TGAGATTACA GGCACTGGCC ACCAGACCCG GCTAATTTTT
12901 TTTTCTTTTT CTTTTTTTTG AGACGGAGTT TCGCTCTTTG TTGCCCAGGC
12951 TGGAGTACAG TGGTGTGATC TCGGCTCACT GCAACCTCCG CCTCCCGGGT
13001 TCAAGTGATT CTCCTGTCTC AGCCTCCCGA GTAGCTGGGA TTACAGGCGC
13051 GCGCCACCAT GCCTGGCTAA TTTTGTATTT TTTTTTTTTT GAGACAGAGT
13101 CTCACTCTGT CACCCAGACT GGAGTGCGGT GGCGCGATCT CGGCTCACTG
13151 CAAGCTCTGC TTCCCGGGTT CATGCCATTC TCCTGCCTCA GCCTCCGGAG
13201 TAGCTGGGAC TACAAGCACC CACCACCGTG CCCGGCTAAT TTTTTGTATT
13251 TTTAGTAGAG ACGGGGTTTC ACCGTGGTCT CGACCTCCAG ACCTCGTGAT
13301 CCACTAGCCT CAGCCTCCCA AAGTGCTGGG ATTACAGGCG TGAGCCACCT
13351 CACCCAGCCT AATTTTGTAT TTTTAGTAGA GATGGGGTTT CACCATGTTG
13401 CGCAGGCTGG TATTGAACTT CTGACCTCAG GTGATCCGCC CGCCTCGGCC
13451 TCCCGAAGTT CTGGGATTAT AGGCGTGAGC CACCGCACCT GGCCTAATTT
13501 TTGTATTTTT AGTAGAGATG GAGTTTTACC TTGTTGGCCA GGCTGGTCTT
13551 GAACTCCTGA CCTCACCTCA GGTGATCTGC CCACCTCGGC CTCCCAAAGT
13601 GCTGGGATTA CAGGCATGAG CCACTGTGCA CCCGGCCTAA AAATCACCAT
13651 CTTGACAGAA CTTCACGCCT TGCTTTTTGT TTTTTTTCAT CTTTGTGCTT
13701 GTTTTCCACT TAACCCTTGA TCACAGACAT CTTTCCATGT GGATTCATGT
13751 AGAACTACCT CATTCGTTAG AACAGCTGCA GAGTATTCCA CTGTGCGGTT
13801 AGTCCATCAT TTCCCTAACC ATCCTCCTGC TGATGGACAG TTAGACTGTT
13851 CCAGTTTTTC AGTATGATTC TATGCCAGGC TGCCATGAAC GTCCTTTTAC
13901 TGATCCACTC AGGCCAGTAT TTCTGTAGGA GAAATTCCTA GAAGTGGGAT
13951 AATTGGATCA AAAGATATGC ACATTCTAAA TTAGGAGAGA GACTGCCAAA
14001 CTGACCTCAG ACAAGGTTGT ACCAGTTTGC ACCCCCATCA GCAGCGTACA
14051 AGTGCCTGCT TCCCAACTTC CTCGCCAACA GGGATGCTAT AAAAAGCTTC
14101 ACAATTTTGC CAGTCTCATT GGCAAATGGT ATCTTGGTTA AATTTGCATT
14151 TCTTTAATAC TAAGTGGGGG TAGGGTATCT TTTCATATGT TTATTGGCCA
14201 TTTATTTCTT CTGTCAATTG CCTGTTCTGA TTCCTTGTCC ATTATTCTAC
14251 TGGGTTTGTT GGTCTTTTTC TCATTGATTT TTAGAATCTC TGTTAATGGA
14301 TATTAACCCT TTGCTGTTGA ATGTGTTTGC AAATATTTTC TCCCTGTCTG
14351 TCATTTATGT GTCTTTTTCC ATATAAATTT AAAAAATTTT GGTGGGCTCA
14401 ATAGGTCAGT CTTTCCCTTC CGGGCTTCTG GGATTTGTGT TCGGGGTAGA
14451 AAGGCCCTCA GCCCCTCAAG ATTATAAAAT TATAAAACCT TTTCTTTTTT
14501 TTTTTTTTTT CTGAGACAGG GTGTCTTGCC ATGTCACCCA GGCTGGAGTG
14551 CAGTGGCATG ATCTTGGCTC GCTGCAACCT CCACCTCCCA GGTTCAAGTG
14601 ATTCTCGTGC CTTAGCCTCC CGAGTAGCTG GGATTATAGG TGCCTGCCAC
14651 TATGCCTGGC TAATTTTTTG TATTTTTAGT AGAGACGGGG CTTTGCCATG
14701 TTGGCCAGGC TGGTCTTGAA CTCCTGACCT CGTGATCCAC CCGCCTTGGC
14751 CTCCCAAAGT GCTGGGACTA CAGGCGTAAG CCACTGTGCT CGGCCCTATA
14801 TTTTTTTCAG ATAGCCAGTT ATCCTAATGC TCCCTTGATT TGATGGACCA
14851 CCTGGATCAC ACATTATGAG CCCCCTCATA AGCAGGTGGG AGTCTCAAGC
14901 GAGGGCCAGT CCCGATGGGA ATAGCACTTG GTGGCTGAGG ACCCTCCTAT
14951 CTGTGCAGAC ACTGTTGTAA AACTTCACAT GCATCATCTA ATTTAGTCCT
15001 CACCAAAATC CTATGAAATG TAGGAATGAT CATTACACCC ATTTATAGAT
15051 AAGGAAACGG AGGGACAGGG AGATTACTCC GCTACAGGTC AAGAGGCAGG
15101 GAAGTAGAGC TGCGATTTGA ACTGAGGTCT GTGTCTAGAA CACGTGCTCA
15151 TTCTTTCCCT AAAATGTATT CATAGGTGAA AAAGGGCTTC TGCGGAAAGC
15201 CCTGGGTTAT GTGGGAAACC CTGGATTTAC AGCTGTCTTT CCAGCAGGAT
15251 GATGCAGGAG AGAGAGGGAT GCGATTTCTC CCAATCTCTC CTGGTCCCAG
15301 AACTCATTAG AGAGTTCTCC CTGCTGAGGG CTCCCGACTG GTGTTGCACA
15351 CAGTACACTT CGGGAGCCCG AGGCTGATGG TTCCATGGAA AGTACACAGT
15401 CATTTTAGTT TGCACACCAA GTGTGAAGTG GGCAGGACAG GCCACTGTTC
15451 TGAGAAGGAA CCCAGGGAAA GGGACTGGCC CAAGACCACA CACTGGTTAG
15501 CGGCACTTCC CACATCTGCC TGACCCCTAG TCCAGTGCCG CCTTTTCTTT
15551 ACTCTGCAAC AGGAGTCCAA AATCAGGAGT TCCATGAGGA CACTGGGAAC
15601 AGTGGGATGG GTTAGGCCAG CGGTGGATGG TTCTGGGGAG GGCCCGAGCT
15651 GAAGCGCCCC CGCAACTCCC CACAGGGATG GCTGCATCAG CAGGGAGGAG
15701 ATGGTTTCCT ATTTCCTGCG CTCCAGCTCT GTGTTGGGGG GGCGCATGGG
```

```
15751 CTTCGTACAC AACTTCCAGG AGAGCAACTC CTTGCGCCCC GTCGCCTGCC
15801 GCCACTGCAA AGCCCTGGTG AGAGTCCCTT TCCCGGCTCA CGGCCCAAGC
15851 CACGCCCCTC CAGCCCCGGC CCCGCCCTCC CTTCTGGCCC CGCCTCTGCC
15901 AGAGCCCTTC TCAAGCCAGG AAAACCTGGT AATTCTATTT GCCTCTCCTC
15951 CTGTGGTTCT GCCCGGGGCC CTGAGGCGGG CTCTAAAGCC CTAGTCTCAC
16001 CCTCAAGAAG GAAGAAGTAG AGTCATCACC TCTAAATCCC TCCTCCCACC
16051 ACGGCCCCTC CTCTATTGCA GATCCTGGGC ATCTACAAGC AGGGCCTCAA
16101 ATGCCGAGGT GAGATGGAAT GACTGGAAGG GCTGCTGGGC AGTGTTTTTT
16151 TTGTTTGTTT GTTTGTTTGG GAGAGTTACT ATTTTGGTGG GGCAATTGCC
16201 AAGGAGTGAA GTACCTTAAA ATCAGAGGCG CATGGCCGGG CATGGTGGCT
16251 CAAGCCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGCGCGC AGATCACCTG
16301 AGGTCAGGAG TTCAAGACCA GCCTGACCAA CATAGCGCAA CCCCGCCTCT
16351 ACTAAAAATA CAAAAGTAG CTGGGCGTGG TGGCACCCAC CTGTAATCCC
16401 AGCTACTTGG GAGGCTGAGG CATGAGAATC GCTTGAACCT GGGAGGCGGG
16451 GTTTGCAGTG AGCCGAGATC ACGCCACTGC ACTCCAGCCT GGGCAACAGA
16501 GAGGGCTCTG TCTCAAAAAA AAAAAACAAC AAAAAAACCC CCAAAACCAA
16551 AACCCCACAA AATCAGAGGC TCAAGATGAC TGATGTGAAG GGAGTGGCGT
16601 TTAAGAGGCC ATTTATTTTG ATGACGCAGC TGCCCAGGAA CAGAGAACAT
16651 GGGAGAAGGC ATAGACTGAC AATTAGGAGG AGGAGAACAC TTTGGAAGGA
16701 GACTCTTATT TTGGTGGGGC AGCTGCTCAG GAACAAAGGT TCCTGGTAGG
16751 GGGGCGCAAG CCTGCGGGAT GGGATGGAGG GTATTCTGAC CAATGTCCCT
16801 GGCTGGCTCT CCATTTGCTC TCCCCCAGCC TGTGGAGTGA ACTGCCACAA
16851 GCAGTGCAAG GATCGCCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG
16901 TGAGCCTGGA GGGGTCTGCA CCCTCACCCT CACCCATGCA CAGCCACCAT
16951 CACCGCGCCT TCAGCTTCTC TCTGCCCCGC CCTGGCAGGC GAGGCTCCAG
17001 GCCTCCAGGT AAGAGGGAGT CATTCTGTAC TGGCCTGTGG AGGGAAGGAT
17051 GCAGGGCTAC TGGGGCAAAG AACGCAGGAT GGAAGCCATT CCAAAGTGCA
17101 TAATTCTCTT TTTGTGGTGG GATAATAAAG AAGGGACAGG CCGGGCGCGG
17151 TGGCTCACGC CTGTAATCCC AGCACTTTGG GAGGCCGAGG CGGGCGGATC
17201 ACGAGGTCAG GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC
17251 TTTACTAAAA ATACAAAAAA AAAAAATTAG CCAGGCGTGG TGGCGGACGC
17301 CTGTAGTCCC AGCTACTTGG GAGGCTGAGG CAGGAGAATG GCATGAACCC
17351 GGGAGGCGGG GCTTGCAGTG AGCCGAGATC GCGCCACTGC ACTCCAGCCT
17401 GGGCGATAGA GCAAGACTCC GTCTCAGAAA AAAAAAAAT AAAAAATAAA
17451 GAAGGGACAG GTAAGGGTGC CAGAAAGTGG CCAGGAAGCC CTGGACCTTC
17501 TGAGGCTGAG GAGAGAGACC CTAATTTATA AAGAGGTATA AAAGTGAAAG
17551 AGGCTTCAAG ATTCCAGTTA CAGTCTTATT TTGTTGGAGG GGTTAACAAA
17601 GGATTGGAGA AGGTGTTATA TGAGCCATTG GCTTGCCTTT CCCTTTCTGG
17651 CTGCTCTGGA GGCTCTTCTG GGGAAAGTCC CTTGCCCTGA TAATGTCCTG
17701 GCAGCTCTCT TGGGGTATTT GATGGTTTTA GGTCAGTTTG CTGAATGACA
17751 ACTGGCCAAA TGATTATTTT GCTGAGAACA GTCCGAACAA CTATGTTAAA
17801 CTGGGGTCTA AGGTAGTTGA TCACAACTGT TTGGGTTGGC ATAAGTCCTC
17851 AAAAACAGA GGCAGGCACA GGGCATACAT CCTCAAAAAT AGAAAGATA
17901 AATCCATTTG CATTGAGCCT TCCAGAAGTG CTGGGGTCTA AAATGTGAAA
17951 TACACACAAA ATTGACATTT AAGCAAACTG CGCTGACAAA TCTGTGGCTG
18001 AAAAGCTGT GGCAAAACAA AAACATAGAA AAAGAGCCTC AAAAATTGGG
18051 CTGAGGCCAG GCATGGTGGC TCACGCCTGT AATCCTAGCA CTTTGGGAAG
18101 CCAAGGTGGG TGGATCACCC GAGGTCAGGA GTTGGAGACC AGACTGGCCA
18151 ACGTGGCAAA ACATCATCTC TACAATACAA AAATACAAAA ATTAGCTGGG
18201 CGTGGTGGCA GGCGCCTGTA ATCCCAGCTA CTTGGGAGGC TGAGGCACGA
18251 GAATCGCTTG AACCTGGGAG GTGGAGGTTG CAGAGAGCCG AGATTGCGCC
18301 ATTGCACTCC AGCCTGGGCG ACAGAGAGAG ACTCTGTCTC AAAAAAAAAA
18351 AAAAAAAAAA AAATTGGGCT GTGAGGTCAT GCAGGGAATT GATTTTTGGT
18401 GGGTGGGTCT GCTTCTGGGA TGATGTGGAT GCCTCCCGTG GAGAGGGGAA
18451 GGGTTGATGA AGTCCCAGGG ACCTGGAAGT GTGTTCTGCA GCAATCCCCC
18501 TCCCAGCAGA GATCCGTGAG GAGGAGGTAC AGACGGTGGA GGATGGGGTG
18551 TTTGACATCC ACTTGTAATA GATGGTGAGT CCTCCCACAG CTGGCACCAG
18601 AGCTCCCCAC TGAGGGCTGG GGGGGAGCTG GGGAGTATCA GGGAAATGGG
18651 TGCTTTATCC AAATGGCTCC AAGCCAGGTG GGCTACTACC TTGTTGTTAG
18701 GGGGGTGTCT TCCTCACAAC CTGTTTTTCT CTTCCCAGCT GTGGTTGGAT
18751 CAAGGACTCA TTCCTGCCTT GGAGAAAATA CTTCAACCAG AGCAGGGAGC
18801 CTGGGGGTGT CGGGGCAGGA GGCTGGGGAT GGGGGTGGGA TATGAGGGTG
18851 GCATGCAGCT GAGGGCAGGG CCAGGGCTGG TGTCCCTAAG GTTGTACAGA
```

```
18901  CTCTTGTGAA  TATTTGTATT  TTCCAGATGG  AATAAAAAGG  CCCGTGTAAT
18951  TAACCTTCAC  CATCAGCGCC  TAGAATCCCG  GGGGGTAGGG  GGATGGTATA
19001  CTTTACAGGA  TGACAATCTT  GGGAGCTAGA  ACTTTGTAGC  CAGAGAAACT
19051  TGGGAGGTCT  GGAATCTCAT  GTGTCTGGAG  TCTTGGGGAA  GAGAATCTTA
19101  GAAGCAGAAA  ACCTTGGAAC  ATAAGAATCT  TGGGGAGGGT  CTAGGATCTT
19151  GAGGAGACCA  GATCCTTGGA  CATCTAAAAC  TTGAAACTAG  TAGGTCTGCA
19201  CCCGAGAATT  GCAGGGCCAG  TCATGCATAC  CCAAAGCCTT  CAGCCCATGG
19251  CCGAAATTCC  CTTGCTGGAC  AGGGGGCCTT  TCAGCCCCTG  CTTGGACGCT
19301  TCCAGTAACA  GGGCCCTCAC  TGCAGGAATC  GTGGGAGGGA  GAGGGGCAGC
19351  ACAGAGTTGC  TGGCTGTCGG  GGAAGGGAGG  GAGGGCCCTG  GGCAGTCCGA
19401  GGGCCCTGCT  GGGCTTGTGC  CTCAGGGTGG  GGGCTGCACT  CCTCCGCCTT
19451  GCAGCCTCCT  GGCCTGGTGC  TGCTGCCAGC  CGGAAGGACA  GTGACTTCCA
19501  GAGGAAATGC  ATATTGATCC  TGCTTTCAGC  CTCCGGTGGT  GGCTTCTCCC
19551  AACCCAGCTC  TTCCCTCCTG  AGCCTGCAGC  ACGGAGGTTT  TGGGGGTCAC
19601  TGCTACCTAA  AGAAGGCTAA  GGCCACTTCT  GAGGCTGGTC  TGGGAGTTTA
19651  CTAAAGGTTC  TGAAGCTGGG  CCGGGCTGCC  CCTGGGATCA  GGAGACTCCA
19701  GACAGCAGTC  CTGACAATGG  GAACTACCTC  CTCAGTCCCC  CAAACTGGGA
19751  GGTGTCCCAC  AGCAGCTGTA  GGATTGTCCT  AGGGGTGGAG  ACCTGAGCAC
19801  CTTCCACTCC  AAAGCACAGT  ATCTGTGGGC  CTGGCAGTGG  CCTCAGTTCC
19851  CCCATGAGTG  CCCCGGTCCC  CCACCCCAGG  GTTTCCCCAC  ATCACATCCA
19901  TCCCTGCTTT  GAGACCCCAC  TCCCCCTGGC  CTGTTCTTTA  TTTTGGGTCA
19951  CTCCCTTCTC  TTTCCTGGTC  ATATCTCTCC  TGCAGGCCTA  CCCTGTGTTG
20001  GGCCCCCCAG  CCCTGTCTCT  GCATCGGGTG  CCCCCCTGCC  CCTCCTTCTG
20051  TCCTCAGCCC  CCTCCGCCCT  TCCCCCTCTT  GAGGCTGTAA  TATCCGTTTC
20101  ACGATTTGGG  GGCTGAGTTG  CTATAACAAC  AGACGGCGAT  TGTGTTGTGA
20151  AGAGCAGCTC  GCTCCTGTGC  CGCCTGCCTC  CTGTGCTGCC  TCCATCCCTG
20201  CAGCCCAGTC  GGTTCCTCTT  GGCTCCTCTC  GTCACTACCC  TCCAGTTCCA
20251  GTCTGGCCTC  TTCCTGGTGT  GTGTGTGTGT  GTGTGTGTGT  GTGTGTGTAT
20301  GCATGCATGC  ATATGTGTGT  CCAGGTCTGC  CTGCCCGGGA  TGTGACAAGT
20351  AGCGGTCTTC  ATGGTTGCAT  GTGTCTGAAT  TTGGTGTCTG  AGCTTCACAT
20401  TGTATGCGCC  TGTGTGCATG  TGTGTGCATG  GACATGCATG  CTGTATCTGC
20451  TGTGTTTCCC  CTCCCCCATG  TGTCCCCACT  GGCCTTTGCA  CATGGGAGAA
20501  GGGCATGTGC  TCAGCATATC  ACTCAACTGT  CCACATTGGG  TGGGTACCTG
20551  TGTGTGGTGT  GTGTGTGTGG  GGGGTGTGTC  TTGAAGTGGC  AGGTCCCAAA
20601  TGCTTAGGCA  ATCTGAACCT  TGGACCTTGC  AGAGAGGAGA  GATGTCCCTG
20651  TAGGTGGGAG  GGACAGGGAG  ATGCAGCAGC  TGCCCGGTGA  CCTTTTCTGC
20701  CCTTGATGGG  CAAAGCTGGG  GGTAGGGAAA  GGAGACAAGT  GCTCATACTT
20751  ACCTCCCTCC  CTGCCCAGGC  TCCTCTGTAA  GGGTCTGAGT  CTGTCTCTGT
20801  GAGCCATTGC  ATCTGTCTGT  CTATGCCCTG  ATGCCTGGAT  GGACAAGGGG
20851  TGTGTGTGTG  TGTGTGTGTG  TGTGTGTGTG  AGTGTGAGGC  TGCAGGAAGA
20901  GGAACAGTGG  GGGATGGGCA  GGAAAGTGGG  CTGTGGGGTC  AGGGAGGCGA
20951  T  (SEQ ID NO:3)
```

FEATURES:

```
Start:     3000
Exon:      3000-3072
Intron:    3073-3753
Exon:      3754-3855
Intron:    3856-4363
Exon:      4364-4427
Intron:    4428-4786
Exon:      4787-4918
Intron:    4919-5702
Exon:      5703-5853
Intron:    5854-6056
Exon:      6057-6230
Intron:    6231-6389
Exon:      6390-6506
Intron:    6507-8832
Exon:      8833-9114
Intron:    9115-9885
Exon:      9886-9963
```

```
Intron:   9964-10201
Exon:     10202-10324
Intron:   10325-10638
Exon:     10639-10754
Intron:   10755-15675
Exon:     15676-15817
Intron:   15818-16071
Exon:     16072-16108
Intron:   16109-16828
Exon:     16829-17008
Intron:   17009-18491
Exon:     18492-18565
Stop:     18566
```

CHROMOSOME MAP POSITION:
Chromosome 11

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain
5539	C	G	Intron
5658	T	G	Intron
5861	C	T	Intron
6023	A	G	Intron
6799	C	T	Intron
9579	C	A	Intron
9842	T	C	Intron
10159	T	C	Intron
12025	A	– G	Intron
14723	T	C	Intron
14996	G	A	Intron
16153	T	G	Intron
16181	G	A	Intron
16756	A	G	Intron
18059	A	G	Intron
18364	A	– T	Intron
18861	G	A	Beyond ORF(3')
20443	G	A	Beyond ORF(3')
20881	A	T	Beyond ORF(3')

Context:

DNA Position	
5539	AGACTCAGTTCAGGCATGAAGTCTCCGTGGGCTCTGAGGGTTCGGGGCTCTTCCGGGGTA
	GAATTTGTCGTTCCCACCTCTGTTTTTCCATGGCACTTTGTACAGACTCCTGTACAAAGAC
	CTCTGTACATGTGTCACGCTGTTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGTA
	GACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCCT
	AGCACAGTGCCTGGCACAGAGTACGTTGTTCATAAATGTGTGTTGAGTGCATGACGGGGT
	[C,G]
	GGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTCGTGCCTAGGACAGTGCCTCGC
	ATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGGCCACGAGGC
	ATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCCCAGCCCTACCTACAAGTGGA
	AGCGGCAGGTGACTCAGCGGAACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGT
	TTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCT
5658	CCTCTGTACATGTGTCACGCTGTTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGT
	AGACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCC
	TAGCACAGTGCCTGGCACAGAGTACGTTGTTCATAAATGTGTGTTGAGTGCATGACGGGG
	TGGGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTCGTGCCTAGGACAGTGCCTC
	GCATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGGCCACGAG

[T,G]
CATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCCCAGCCCTACCTACAAGTGG
AAGCGGCAGGTGACTCAGCGGAACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTG
TTTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCC
TTCTGCAAGATCCTGGTGCGGCCCGAGGGCTGGGGGGTCAGGGGTCCAATGTGGGCTGGA
AGAGAGTTCTAGGAGGGGCAGGGTCCCTGGCGTAGGCTGGGTCACAGGGTGCATCAGGGG

5861   CTGGGACTGGGAACATTCGTGCCTAGGACAGTGCCTCGCATTATGTAGGTTCTCAGTAAG
        CGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAGGCATGCGCATGTCCAGCAAAGGG
        CTCACTACCCCTGCCCCCCCAGCCCTACCTACAAGTGGAAGCGGCAGGTGACTCAGCGGA
        ACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGTTTGACCACCTGGAGCCCATGG
        AGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCTTCTGCAAGATCCTGGTGCGGC
        [C,T]
        CGAGGGCTGGGGGGTCAGGGGTCCAATGTGGGCTGGAAGAGAGTTCTAGGAGGGGCAGGG
        TCCCTGGCGTAGGCTGGGTCACAGGGTGCATCAGGGGTTTCAGTGTAACCACTGAAGGTC
        AGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGGAGAGGCCGGCAAGGTGCTGAGGCCAC
        TCCTCATGCCCCCAGTTTCAGGACTATCACAGTTTCGTGACTCATGGCTGCACTGTGGAC
        AACCCCGTCCTGGAGCGGTTCATCTCCCTCTTCAACAGCGTCTCACAGTGGGTGCAGCTC

6023   GGCAGGTGACTCAGCGGAACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGTTTG
        ACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCTTCT
        GCAAGATCCTGGTGCGGCCCGAGGGCTGGGGGGTCAGGGGTCCAATGTGGGCTGGAAGAG
        AGTTCTAGGAGGGGCAGGGTCCCTGGCGTAGGCTGGGTCACAGGGTGCATCAGGGGTTTC
        AGTGTAACCACTGAAGGTCAGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGGAGAGGCC
        [A,G]
        GCAAGGTGCTGAGGCCACTCCTCATGCCCCCAGTTTCAGGACTATCACAGTTTCGTGACT
        CATGGCTGCACTGTGGACAACCCCGTCCTGGAGCGGTTCATCTCCCTCTTCAACAGCGTC
        TCACAGTGGGTGCAGCTCATGATCCTCAGCAAACCCACAGCCCCGCAGCGGGCCCTGGTC
        ATCACACACTTTGTCCACGTGGCGGAGGTGCCTGCCCCTCCCTCCCGGTGTCTCCCAACC
        ACCCCACATGCCAGTCAGGCCAACCCTTCCCTTCCCCTAACCCACTGCCTTCTCTCTAGA

6799   CCATCAAGGTGCCTGGGACTGGGGAGGGGCCGGTGCTTCCCAGGTCTGTCTTCACTGGGT
        CCTCCCAGCAGCACTGGGGGCTGGGCACAGCTGTCCTCATTTGATAGATATGGAAATGGA
        GGCTCAGAGGGGTTAAGTGCTTTTCTCAGTTTGCACAATGGCAACAGCAGAGTGGGGGCT
        CACAGGTCGTCAGGGACCCCAAAGCTAGTACTTTTTTTTTTTTTTTTAAGACAGGGTCTC
        TCTCTCTGTTGTCCAGACTGGAGTTCAGTGGTGCAGTCACAAGCTCACTGCAGCCTTGAA
        [C,T]
        TCCTGAGCTCAATCGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGTAC
        GCCACCATGCCTAATTTTTTGTATTGTTATTAATTTTTTTTTTTTTTTTTAGAGATGGGG
        TTTTGCCATGTTGCCCAGACTGGTCTTGAACTCCTGGGCTCAAGTGATCCGCCTGCCTTG
        GCCTCCCAAAGTGCTGAGATTATGGCTTGAGCCATTGTGCCTTGCCACTTGTAGTTTCTT
        CTTTTCTTTCTCCTTCATTTTTTATTATTTTTGAAGTATTTTGAAGTATTGAGTAACATA

9579   CTGGTTGGGGCTGTGGACTGAGGTCTGATCTCCAGGCTGGTATGTGGACTGTGGGCAGTT
        TGAACTGGGCCTGGGTCCCGGGTTGAGTTCTGGCAATGGGCTGTGTTCTAGGGCTGGGCC
        AAGCTCTGCATTCTGTGGGCAGGGGTGGTTTCTAAGCATGGCCCTGGGCTCGGAGTGAAG
        TTCTGGGCTTGGCTTTACACTTGGTCTTGGGGTCTAGGGTGGGAGTTGGGTTCTGGTTTA
        GATCCAGACAAGGTTCTAGACATTGGGCTGGGGCTTAAGTGTTAAGGTTTGGAGTGGATT
        [C,A]
        TTAGCTGCTTCTGGGCTCTGGAGGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTC
        CGACCTGGCTTCTTCCCTGACATCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGC
        TGTGGTTTGATCTGTGCGCTGGGATGACATGGGGGTTGCTGTGCTGTGTTCTAAGCCAGG
        CTTTGTCCTGAGTCTAGCTTCTGACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTCGAC
        CTTTGGCCCTGGGCTCTGTGGCCGTGGGCAGGGGCCAGTGGGGGGTGATCAGATCTGTGTG

9842   TGGGCTGGGGCTTAAGTGTTAAGGTTTGGAGTGGATTCTTAGCTGCTTCTGGGCTCTGGA
        GGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTCCGACCTGGCTTCTTCCCTGACA
        TCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGCTGTGGTTTGATCTGTGCGCTGG
        GATGACATGGGGGTTGCTGTGCTGTGTTCTAAGCCAGGCTTTGTCCTGAGTCTAGCTTCT
        GACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTCGACCTTTGGCCCTGGGCTCTGTGGC
        [T,C]
        GTGGGCAGGGGCCAGTGGGGGTGATCAGATCTGTGTGTCCCAGGTGTCTCTGGATCAGTA
        TCAGACGGAGGATGAGCTGTACCAGCTGTCCCTGCAGCGGGAGCCGCGCTCCAAGTCCTC

```
            GGTGAGGGGGTACTCCCTCCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTG
            AGGATGGGAAGAGCTCTTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCC
            TGGGGGTAGGGCAGTAGTGTTGGGCAGACTTCCCTCTCCCAGGGATTCCCCTCTCTGTTC

   10159    GGGGGTGATCAGATCTGTGTGTCCCAGGTGTCTCTGGATCAGTATCAGACGGAGGATGAG
            CTGTACCAGCTGTCCCTGCAGCGGGAGCCGCGCTCCAAGTCCTCGGTGAGGGGGTACTCC
            CTCCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTGAGGATGGGAAGAGCTC
            TTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCCTGGGGGTAGGGCAGTA
            GTGTTGGGCAGACTTCCCTCTCCCAGGGATTCCCCTCTCTGTTCCCCGGGGCTCTGGGCT
            [T,C]
            CCCCTGCCTCTGGCCCTAGCTCAGGCCCGACCATTTCCATAGCCAACCAGCCCCACGAGT
            TGCACCCCACCACCCGGCCCCGGTACTGGAGGAGTGGACCTCGGCTGCCAAACCCAAG
            CTGGATCAGGCCCTCGTGGTGGAGCACATCGAGAAGATGGTGGAGGTGAGCTCCTGCGGA
            GCCTGAGCAGTGTGTGGGGAGAGGCCAGTTTGCCGGAGCACTGCCCTGGAAGCCAGCACG
            AGTGTCCTGTTCAAGACCCAGCACTCAGCCCCTAGGAGTCACAGGGCCTGGCAGGCCAGC

   12025    TGGGGTTTTACCATGTTGGCCAGGCTGGTCTCAAACTCTTGACCTCAAGTGATCTACCTG
            CCTTGGCCCCCCAAAGTGCTAGGATTACAGGCATGAGCTACTGCTCCTAGCCTAAAAAAA
            TTTTTTTTGGGCATGGGTGGCACGTGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAG
            GAGGAACCCTTGAGCCCAGGAGGTTGAGACTGCAGTGAGCTGTCATCACACCACTGCACT
            TCAGCCTGGGTGACTGCGCGAGATCACCCCCATCAAAAAAAAAAAAAAAAAAGAAAAAAAA
            [A,-,G]
            GAAGAAATGAAAGTCCCCTCTTTCCTTTTCCACTGGTAGAAGTTGCCATGATTAAGCACT
            GTTAACAATATTAAGCTTGGCAGTATGTGGATTCTTCCAGTCTTCTTTTCCCAGGCAGGT
            GCACATTGATAGAGATTTTGTTTGTTTGGTGTCTGTTTCATGGACAAACAGGATTAGAGC
            ATAAATCTAGTTCTGCTTGTGGCTTTTATCATAGCTGCTTTATTTCTTCTCCCAGATTTT
            AGGCAGAGGTAGTTGAGTTCCATGTTTTCTCCCTGGGTTGGTGGGTGGATTTTTATCTAG

   14723    GGCTTCTGGGATTTGTGTTCGGGGTAGAAAGGCCCTCAGCCCCTCAAGATTATAAAATTA
            TAAAACCTTTTCTTTTTTTTTTTTTTTTTCTGAGACAGGGTGTCTTGCCATGTCACCCAGG
            CTGGAGTGCAGTGGCATGATCTTGGCTCGCTGCAACCTCCACCTCCCAGGTTCAAGTGAT
            TCTCGTGCCTTAGCCTCCCGAGTAGCTGGGATTATAGGTGCCTGCCACTATGCCTGGCTA
            ATTTTTTGTATTTTTAGTAGAGACGGGGCTTTGCCATGTTGGCCAGGCTGGTCTTGAACT
            [T,C]
            CTGACCTCGTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGACTACAGGCGTAAGCCA
            CTGTGCTCGGCCCTATATTTTTTCAGATAGCCAGTTATCCTAATGCTCCCTTGATTTGA
            TGGACCACCTGGATCACACATTATGAGCCCCCTCATAAGCAGGTGGGAGTCTCAAGCGAG
            GGCCAGTCCCGATGGGAATAGCACTTGGTGGCTGAGGACCCTCCTATCTGTGCAGACACT
            GTTGTAAAACTTCACATGCATCATCTAATTTAGTCCTCACCAAAATCCTATGAAATGTAG

   14996    CCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCGTGATCCACCCGCCTTGGCCTCCC
            AAAGTGCTGGGACTACAGGCGTAAGCCACTGTGCTCGGCCCTATATTTTTTTCAGATAGC
            CAGTTATCCTAATGCTCCCTTGATTTGATGGACCACCTGGATCACACATTATGAGCCCCC
            TCATAAGCAGGTGGGAGTCTCAAGCGAGGGCCAGTCCCGATGGGAATAGCACTTGGTGGC
            TGAGGACCCTCCTATCTGTGCAGACACTGTTGTAAAACTTCACATGCATCATCTAATTTA
            [G,A]
            TCCTCACCAAAATCCTATGAAATGTAGGAATGATCATTACACCCATTTATAGATAAGGAA
            ACGGAGGGACAGGGAGATTACTCCGCTACAGGTCAAGAGGCAGGGAAGTAGAGCTGCGAT
            TTGAACTGAGGTCTGTGTCTAGAACACGTGCTCATTCTTTCCCTAAAATGTATTCATAGG
            TGAAAAAGGGCTTCTGCGGAAAGCCCTGGGTTATGTGGGAAACCCTGGATTTACAGCTGT
            CTTTCCAGCAGGATGATGCAGGAGAGAGAGGGATGCGATTTCTCCCAATCTCTCCTGGTC

   16153    CGCCCCTCCAGCCCCGGCCCCGCCCTCCCTTCTGGCCCCGCCTCTGCCAGAGCCCTTCTC
            AAGCCAGGAAAACCTGGTAATTCTATTTGCCTCTCCTCCTGTGGTTCTGCCCGGGGCCCT
            GAGGCGGGCTCTAAAGCCCTAGTCTCACCCTCAAGAAGGAAGAAGTAGAGTCATCACCTC
            TAAATCCCTCCTCCCACCACGGCCCCTCCTCTATTGCAGATCCTGGGCATCTACAAGCAG
            GGCCTCAAATGCCGAGGTGAGATGGAATGACTGGAAGGGCTGCTGGGCAGTGTTTTTTTT
            [T,G]
            TTTGTTTGTTTGTTTGGGAGAGTTACTATTTTGGTGGGGCAATTGCCAAGGAGTGAAGTA
            CCTTAAAATCAGAGGCGCATGGCCGGGCATGGTGGCTCAAGCCTGTAATCCCAGCACTTT
            GGGAGGCCGAGGCGCGCAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGACCAACAT
            AGCGCAACCCCGCCTCTACTAAAAATACAAAAAGTAGCTGGGCGTGGTGGCACCCACCTG
            TAATCCCAGCTACTTGGGAGGCTGAGGCATGAGAATCGCTTGAACCTGGGAGGCGGGGTT
```

<div align="center">

FIGURE 3, page 10 of 12

</div>

```
16181    CTTCTGGCCCCGCCTCTGCCAGAGCCCTTCTCAAGCCAGGAAAACCTGGTAATTCTATTT
         GCCTCTCCTCCTGTGGTTCTGCCCGGGGCCCTGAGGCGGGCTCTAAAGCCCTAGTCTCAC
         CCTCAAGAAGGAAGAAGTAGAGTCATCACCTCTAAATCCCTCCTCCCACCACGGCCCCTC
         CTCTATTGCAGATCCTGGGCATCTACAAGCAGGGCCTCAAATGCCGAGGTGAGATGGAAT
         GACTGGAAGGGCTGCTGGGCAGTGTTTTTTTTGTTTGTTTGTTTGTTTGGGAGAGTTACT
         [G,A]
         TTTTGGTGGGGCAATTGCCAAGGAGTGAAGTACCTTAAAATCAGAGGCGCATGGCCGGGC
         ATGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGCGCAGATCACCTGA
         GGTCAGGAGTTCAAGACCAGCCTGACCAACATAGCGCAACCCCGCCTCTACTAAAAATAC
         AAAAAGTAGCTGGGCGTGGTGGCACCCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGC
         ATGAGAATCGCTTGAACCTGGGAGGCGGGGTTTGCAGTGAGCCGAGATCACGCCACTGCA

16756    CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGAGGGCTCTGTCTCA
         AAAAAAAAAAACAACAAAAAAAACCCCCAAAACCAAAACCCCACAAAATCAGAGGCTCAAG
         ATGACTGATGTGAAGGGAGTGGCGTTTAAGAGGCCATTTATTTTGATGACGCAGCTGCCC
         AGGAACAGAGAACATGGGAGAAGGCATAGACTGACAATTAGGAGGAGGAGAACACTTTGG
         AAGGAGACTCTTATTTTGGTGGGGCAGCTGCTCAGGAACAAAGGTTCCTGGTAGGGGGGC
         [A,G]
         CAAGCCTGCGGGATGGGATGGAGGGTATTCTGACCAATGTCCCTGGCTGGCTCTCCATTT
         GCTCTCCCCAGCCTGTGGAGTGAACTGCCACAAGCAGTGCAAGGATCGCCTGTCAGTTG
         AGTGTCGGCGCAGGGCCCAGAGTGTGAGCCTGGAGGGGTCTGCACCCTCACCCTCACCCA
         TGCACAGCCACCATCACCGCGCCTTCAGCTTCTCTCTGCCCCGCCCTGGCAGGCGAGGCT
         CCAGGCCTCCAGGTAAGAGGGAGTCATTCTGTACTGGCCTGTGGAGGGAAGGATGCAGGG

18059    AATGATTATTTTGCTGAGAACAGTCCGAACAACTATGTTAAACTGGGGTCTAAGGTAGTT
         GATCACAACTGTTTGGGTTGGCATAAGTCCTCAAAAAACAGAGGCAGGCACAGGGCATAC
         ATCCTCAAAAATAGAAAAGATAAATCCATTTGCATTGAGCCTTCCAGAAGTGCTGGGGGTC
         TAAAATGTGAAATACACACAAAATTGACATTTAAGCAAACTGCGCTGACAAATCTGTGGC
         TGAAAAAGCTGTGGCAAAACAAAAACATAGAAAAAGAGCCTCAAAAATTGGGCTGAGGCC
         [A,G]
         GGCATGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACC
         CGAGGTCAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACA
         AAAATACAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGG
         CTGAGGCACGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGC
         CATTGCACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAA

18364    TGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACCCGAG
         GTCAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACAAAAA
         TACAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGA
         GGCACGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGCCATT
         GCACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAA
         [A,-,T]
         TGGGCTGTGAGGTCATGCAGGGAATTGATTTTTGGTGGGTGGGTCTGCTTCTGGGATGAT
         GTGGATGCCTCCCGTGGAGAGGGGAAGGGTTGATGAAGTCCCAGGGACCTGGAAGTGTGT
         TCTGCAGCAATCCCCCTCCCAGCAGAGATCCGTGAGGAGGAGGTACAGACGGTGGAGGAT
         GGGGTGTTTGACATCCACTTGTAATAGATGGTGAGTCCTCCCACAGCTGGCACCAGAGCT
         CCCCACTGAGGGCTGGGGGGGAGCTGGGGAGTATCAGGGAAATGGGTGCTTTATCCAAAT

18861    ACTTGTAATAGATGGTGAGTCCTCCCACAGCTGGCACCAGAGCTCCCCACTGAGGGCTGG
         GGGGGAGCTGGGGAGTATCAGGGAAATGGGTGCTTTATCCAAATGGCTCCAAGCCAGGTG
         GGCTACTACCTTGTTGTTAGGGGGGTGTCTTCCTCACAACCTGTTTTTCTCTTCCCAGCT
         GTGGTTGGATCAAGGACTCATTCCTGCCTTGGAGAAAATACTTCAACCAGAGCAGGGAGC
         CTGGGGGTGTCGGGGCAGGAGGCTGGGGATGGGGGTGGGATATGAGGGTGGCATGCAGCT
         [G,A]
         AGGGCAGGGCCAGGGCTGGTGTCCCTAAGGTTGTACAGACTCTTGTGAATATTTGTATTT
         TCCAGATGGAATAAAAAGGCCCGTGTAATTAACCTTCACCATCAGCGCCTAGAATCCCGG
         GGGGTAGGGGGATGGTATACTTTACAGGATGACAATCTTGGGAGCTAGAACTTTGTAGCC
         AGAGAAACTTGGGAGGTCTGGAATCTCATGTGTCTGGAGTCTTGGGGAAGAGAATCTTAG
         AAGCAGAAAACCTTGGAACATAAGAATCTTGGGGAGGGTCTAGGATCTTGAGGAGACCAG

20443    TGTTGTGAAGAGCAGCTCGCTCCTGTGCCGCCTGCCTCCTGTGCTGCCTCCATCCCTGCA
         GCCCAGTCGGTTCCTCTTGGCTCCTCTCGTCACTACCCTCCAGTTCCAGTCTGGCCTCTT
```

FIGURE 3, page 11 of 12

```
CCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGCATGCATGCATATGTGTGTCC
AGGTCTGCCTGCCCGGGATGTGACAAGTAGCGGTCTTCATGGTTGCATGTGTCTGAATTT
GGTGTCTGAGCTTCACATTGTATGCGCCTGTGTGCATGTGTGTGCATGGACATGCATGCT
[G,A]
TATCTGCTGTGTTTCCCCTCCCCCATGTGTCCCCACTGGCCTTTGCACATGGGAGAAGGG
CATGTGCTCAGCATATCACTCAACTGTCCACATTGGGTGGGTACCTGTGTGTGGTGTGTG
TGTGTGGGGGGTGTGTCTTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGG
ACCTTGCAGAGAGGAGAGATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGC
CCGGTGACCTTTTCTGCCCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCT
```

20881
```
TTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGGACCTTGCAGAGAGGAGA
GATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGCCCGGTGACCTTTTCTGC
CCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCTCATACTTACCTCCCTCC
CTGCCCAGGCTCCTCTGTAAGGGTCTGAGTCTGTCTCTGTGAGCCATTGCATCTGTCTGT
CTATGCCCTGATGCCTGGATGGACAAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
[A,T]
GTGTGAGGCTGCAGGAAGAGGAACAGTGGGGGATGGGCAGGAAAGTGGGCTGTGGGGTCA
GGGAGGCGAT
```